

2 Müntz-Jackson Theorems

Recall:

- **Density properties of Müntz polynomials.**

Theorem (Theorem 1.1 in [Lorentz (1996)]).

Let $\Lambda_\infty = \{0 = \lambda_0 < \lambda_1 < \dots < \lambda_n < \infty\}$ with $\lambda_n \rightarrow \infty$. Then the Müntz space $\mathcal{M}(\Lambda_\infty)$ is dense in each of the spaces $C[0, 1]$ or $L_p[0, 1]$, $1 \leq p < \infty$ if and only if

$$\sum_{k=1}^{\infty} \frac{1}{\lambda_k} = \infty.$$

The density property can be indeed extended in several ways: unsorted sequences (which may result in distinct cluster points), complex sequences, and intervals away from the origin.

Under this iff condition, the sequence Λ_∞ can be generalized as

$$\begin{cases} \lambda_0 \geq 0 \text{ and } \{\lambda_k\}_{k=1}^{\infty} \text{ with } \inf_{k \geq 1} \{\lambda_k\} > 0, & \text{for } L_p[0, 1] \text{ with } 1 \leq p < \infty; \\ \lambda_0 = 0 \text{ and } \{\lambda_k\}_{k=1}^{\infty} \text{ with } \inf_{k \geq 1} \{\lambda_k\} > 0, & \text{for } C[0, 1] \text{ with } p = \infty. \end{cases}$$

But its proof requires some techniques and discussions in the cases of distinct cluster points (see [Borwein (1995), Sec. 4.2] or [Almira (2007), Sec. 3.1]).

- **L_p -Best Approximation by Müntz Polynomials.** Let $f \in L_p[0, 1]$ if $1 \leq p < \infty$ (or $C[0, 1]$ if $p = \infty$). The error of approximation from $\mathcal{M}(\Lambda_n)$ to f is

$$E(f, \Lambda_n)_p := \inf_{M \in \mathcal{M}(\Lambda_n)} \|f - M\|_{L_p[0, 1]}.$$

Note that $C[0, 1]$ with L_∞ -norm is completed since it reduces to **maximum norm**. In fact, when $p = \infty$, for $f \in C[0, 1]$ we have

$$\|f\|_\infty := \inf\{C : |f(x)| \leq C \text{ a.e. on } [0, 1]\} = \inf_{\substack{mF_0=0 \\ F_0 \subset [0, 1]}} \left\{ \sup_{x \in [0, 1] \setminus F_0} |f(x)| \right\} = \max_{0 \leq x \leq 1} \{|f(x)|\},$$

where $mF_0 = 0$ denotes that the Lebesgue measure of F_0 is 0.

Schedule:

We consider the L_p best approximation (or Jackson Theorems in Sec. 2) in several subsections:

1. Existence and uniqueness of L_p best approximation.
2. Error of approximation for monomial x^r , and dense properties.
3. Error of approximation for $f \in W_p^1[0, 1]$, and some corollaries.

Notation Convention:

- **AuxThm** refers to the auxiliary theorem, which is not included in this book, similar to **AuxCor**, **AuxLem**, and other related terms.
- Denote $\Lambda_\infty = \{0 = \lambda_0 < \lambda_1 < \dots < \lambda_n < \dots\}$ with $\lim_{n \rightarrow \infty} \lambda_n = \infty$.
- Denote $\Lambda_n = \{0 = \lambda_0 < \lambda_1 < \dots < \lambda_n\}$ simply by Λ , where the integer $n \geq 1$ is fixed.
- Denote the linear space $\mathcal{M}(\Lambda_n) = \text{span}\{x^{\lambda_0}, \dots, x^{\lambda_n}\}$, associated to Λ_n , with respect to the field of real numbers \mathbb{R} .
- $E(f, \Lambda_n)_p = \inf_{M \in \mathcal{M}(\Lambda_n)} \|f - M\|_p$, where $\|\cdot\|_p$ stands for the $L_p[0, 1]$ norm for $1 \leq p \leq \infty$.

2.1 Existence and uniqueness of L_p -best approximation.

Let $(X, \|\cdot\|)$ be a Banach space with **real or complex** scalars, and $X_n \subset X$ be its finite dimensional linear subspace. The *best approximation* to $f \in X$ from X_n is defined as

$$E(f) := \inf_{p \in X_n} \|f - p\|.$$

AuxThm 2.1 (Theorem 1.1, p.59, [Lorentz (1993)]). *For each $f \in X$, there exists a best approximation to f from X_n .*

Proof. Let $F(p) = \|f - p\|$, $\forall p \in X_n$. Let the **closed and bounded** set $C = \{p \in X_n : F(p) \leq \|f\|\}$. Then that $F(p)$ attains its minimum over X_n is equivalent to attain the minimum over C , that is

$$\inf_{p \in X_n} \|f - p\| = \inf_{p \in C} \|f - p\|.$$

Thus the existence is obvious since C is **compact** and $F(p)$ is **continuous**. □

AuxThm 2.2. *If X is strictly convex, which is characterized by*

$$\begin{cases} \forall f_1 \neq f_2, & \|f_1\| = \|f_2\| = 1, & \alpha_1, \alpha_2 > 0, & \alpha_1 + \alpha_2 = 1, \\ \text{imply} & & & \|\alpha_1 f_1 + \alpha_2 f_2\| < 1. \end{cases}$$

*Then the best approximation to $f \in X$ from X_n is **unique**.*

Proof. Suppose that there are $p_1, p_2 \in X_n$ such that $\|f - p_1\| = \|f - p_2\| = E(f)$. If $E(f) = 0$, then $\|p_1 - p_2\| \leq \|f - p_1\| + \|f - p_2\| = 2E(f) = 0$, which implies $p_1 = p_2$. If $E(f) > 0$, we prove it by supposing that $p_1 \neq p_2$, which leads to a contradiction:

$$E(f) \leq \left\| f - \frac{1}{2}(p_1 + p_2) \right\| = \left\| \frac{1}{2}(f - p_1) + \frac{1}{2}(f - p_2) \right\| < E(f).$$

□

AuxLem 2.1. $L_p[a, b]$ is strictly convex for $1 < p < \infty$.

Proof. For any $f_1 \neq f_2$, $\|f_1\|_p = \|f_2\|_p = 1$, $\alpha_1, \alpha_2 > 0$, $\alpha_1 + \alpha_2 = 1$, then by Minkowski's inequality (or triangle inequality):

$$\|\alpha_1 f_1 + \alpha_2 f_2\|_p < \alpha_1 \|f_1\|_p + \alpha_2 \|f_2\|_p = 1.$$

The equality for $1 < p < \infty$ if and only if f_1 and f_2 are **positively linearly dependent**, that is, $f_1 = \lambda f_2$ for some $\lambda \geq 0$ or $f_2 = 0$. This is impossible since $f_1 \neq f_2$ and $\|f_1\| = \|f_2\| = 1$. □

Remark 2.1. Both $L_1[a, b]$ and $L_\infty[a, b]$ are **not** strictly convex.

Remark 2.2. When we consider vectors in \mathbb{R}^2 , the strictly convex property for L_p is **visualizable**. Let $\mathbf{x} = [x_1, x_2]$.

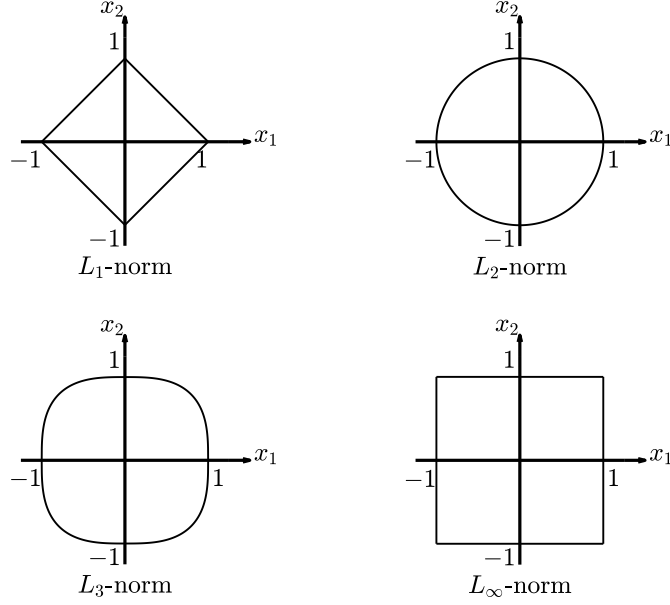


Figure 1: Unit circles in \mathbb{R}^2 with different L_p -norms. Download the [Figure](#) and [Code](#).

AuxThm 2.3. Let $X = C[a, b]$ and $X_n \subset X$ satisfy the **Haar condition**:

For any basis $\{\phi_i(x)\}_{i=1}^n$ of X_n and any set of distinct points $\{\xi_i\}_{i=1}^n \subset [a, b]$, it follows that

$$\begin{bmatrix} \phi_1(\xi_1) & \cdots & \phi_1(\xi_n) \\ \vdots & & \vdots \\ \phi_n(\xi_1) & \cdots & \phi_n(\xi_n) \end{bmatrix} \text{ is non-singular.}$$

Then for any $f \in X$, there is just one L_1 (or L_∞) best approximation to f from X_n .

Proof. The L_1 -best approximation, see [Powell (1981), Theorem 14.3, p.170], while L_∞ -best approximation, see [Powell (1981), Theorem 7.6, p.80]. \square

2.2 Error of approximation for monomial x^r .

Schedule of this Subsection:

- Prove that $E(x^r, \Lambda)_2$ (Eq. (2.1)) and $\mathcal{M}(\Lambda_\infty)$ is dense in $L_2[0, 1]$;
- Prove that $E(x^r, \Lambda)_\infty$ (Eq. (2.2)) and $\mathcal{M}(\Lambda_\infty)$ is dense in $C[0, 1]$;
- Prove that $E(x^r, \Lambda)_p$ ($2 < p < \infty$) (Theorem 2.2) and $\mathcal{M}(\Lambda_\infty)$ is dense in $L_p[0, 1]$.

2.2.1 Case 1: $p = 2$.

Our goal is to prove (2.1) in [Lorentz (1996)], which is stated as following theorem:

AuxThm 2.4 (see also Theorem 5.4 in [Lorentz (1993)]). For $r > -1/2$, $\Lambda = \{\lambda_0, \lambda_1, \dots, \lambda_n\}$ with distinct elements and $\lambda_k > -1/2$, $k = 0, 1, \dots, n$, we have

$$E(x^r, \Lambda)_2 = \frac{1}{\sqrt{2r+1}} \prod_{k=0}^n \frac{|r - \lambda_k|}{|r + \lambda_k + 1|}.$$

Preliminaries.

In a **real** Hilbert space $(H, (\cdot, \cdot))$ with its norm induced by $\|f\| = \sqrt{(f, f)}$, let $f_1, \dots, f_n \in H$ be linearly independent elements, and let $X_n := \text{span}\{f_1, \dots, f_n\}$.

AuxThm 2.5. *For $g \in H$, there is a **unique** $f \in X_n$ such that*

$$\|g - f\| = \inf_{p \in X_n} \|g - p\|.$$

Proof. Existence is obvious and uniqueness follows from that Hilbert space is strictly convex, see [details](#). \square

We call f the *best approximation* of g from X_n in H .

AuxCor 2.1. *Let f be the best approximation of g , then it is equivalent to the orthogonal projection: $(g - f, p) = 0, \forall p \in X_n$.*

Proof. (\Leftarrow) Let f be the orthogonal projection of g onto X_n , i.e., $(g - f, p) = 0, \forall p \in X_n$. Then

$$\|g - p\|^2 = \|g - f + f - p\|^2 = \|g - f\|^2 + \|f - p\|^2 \geq \|g - f\|^2.$$

Thus f is the best approximation.

(\Rightarrow) Let f satisfy $\|g - f\| = \inf_{p \in X_n} \|g - p\|$. For any $p \in X_n$, let

$$h(t) = \|g - (f + tp)\|^2 = \|g - f\|^2 - 2t(g - f, p) + t^2\|p\|^2.$$

Since $h(t)$ achieves its minimum at $t = 0$, then $h'(0) = 0$, which leads to $(g - f, p) = 0$. \square

AuxLem 2.2. *The distance of best approximation $d := \inf_{p \in X_n} \|g - p\|$ is given by*

$$d^2 = \frac{G(g, f_1, \dots, f_n)}{G(f_1, \dots, f_n)},$$

where G is the Gram determinant

$$G(f_1, \dots, f_n) = \begin{vmatrix} (f_1, f_1) & \cdots & (f_1, f_n) \\ \vdots & & \vdots \\ (f_n, f_1) & \cdots & (f_n, f_n) \end{vmatrix}.$$

Proof. The best approximation $f \in X_n$ to g satisfies $(g - f, p) = 0, \forall p \in X_n$. Now we suppose that $f = \sum_{i=1}^n a_i f_i$, then

$$\sum_{i=1}^n a_i (f_i, f_k) = (g, f_k), \quad k = 1, 2, \dots, n. \quad (1)$$

On the other hand, since $(g - f, f) = 0$, $d^2 = (g - f, g - f) = (g, g - f) = (g, g) - (g, f)$, we have

$$\sum_{i=1}^n a_i (g, f_i) = (g, g) - d^2. \quad (2)$$

Hence combining (1) with (2) we have

$$\begin{bmatrix} 1 & (g, f_1) & \cdots & (g, f_n) \\ 0 & (f_1, f_1) & \cdots & (f_n, f_1) \\ \vdots & \vdots & & \vdots \\ 0 & (f_1, f_n) & \cdots & (f_n, f_n) \end{bmatrix} \begin{bmatrix} d^2 \\ a_1 \\ \vdots \\ a_n \end{bmatrix} = \begin{bmatrix} (g, g) \\ (g, f_1) \\ \vdots \\ (g, f_n) \end{bmatrix},$$

and by Cramer's rule,

$$d^2 = \frac{G(g, f_1, \dots, f_n)}{G(f_1, \dots, f_n)}.$$

\square

Remark 2.3. $G(f_1, \dots, f_n) \neq 0$ if and only if f_1, \dots, f_n are linearly independent.

Remark 2.4. *AuxThm 2.5*, *AuxCor 2.1*, and *AuxLem 2.2* provide a **general framework** to compute error estimation of best approximation in a Hilbert space.

AuxLem 2.3 (Cauchy's determinant). For real numbers a_i and b_k that satisfy $a_i + b_k \neq 0$, $1 \leq i, k \leq n$, we have

$$\begin{vmatrix} \frac{1}{a_1+b_1} & \cdots & \frac{1}{a_1+b_n} \\ \vdots & & \vdots \\ \frac{1}{a_n+b_1} & \cdots & \frac{1}{a_n+b_n} \end{vmatrix} = \frac{\prod_{n \geq i > k \geq 1} (a_i - a_k)(b_i - b_k)}{\prod_{1 \leq i, k \leq n} (a_i + b_k)}.$$

Proof. We denote $D(n) = \det[1/(a_i + b_k)]_{1 \leq i, k \leq n}$. We **subtract the last row of $D(n)$ from each of the other rows**, and then we can factor out from $D(n)$ by $1, \dots, n-1$ rows and $1, \dots, n$ columns

$$D(n) = \begin{vmatrix} \frac{1}{a_1+b_1} & \cdots & \frac{1}{a_1+b_n} \\ \vdots & & \vdots \\ \frac{1}{a_{n-1}+b_1} & \cdots & \frac{1}{a_{n-1}+b_n} \\ 1 & \cdots & 1 \end{vmatrix} \cdot \frac{\prod_{i=1}^{n-1} (a_n - a_i)}{\prod_{i=1}^n (a_n + b_k)}.$$

Next we **subtract the last column from each of the other columns**, and extract the factors by $1, \dots, n-1$ rows and $1, \dots, n-1$ columns

$$D(n) = \begin{vmatrix} \frac{1}{a_1+b_1} & \cdots & \frac{1}{a_1+b_{n-1}} & \frac{1}{a_1+b_n} \\ \vdots & & \vdots & \vdots \\ \frac{1}{a_{n-1}+b_1} & \cdots & \frac{1}{a_{n-1}+b_{n-1}} & \frac{1}{a_{n-1}+b_n} \\ 0 & \cdots & 0 & 1 \end{vmatrix} \cdot \frac{\prod_{i=1}^{n-1} (a_n - a_i)}{\prod_{i=1}^n (a_n + b_k)} \cdot \frac{\prod_{k=1}^{n-1} (b_n - b_k)}{\prod_{i=1}^{n-1} (a_i + b_n)}.$$

Therefore

$$D(n) = D(n-1) \cdot \frac{\prod_{k=1}^{n-1} (a_n - a_k)(b_n - b_k)}{(a_n + b_n) \prod_{k=1}^{n-1} (a_k + b_n)(a_n + b_k)},$$

by the induction we complete the proof. □

Proof of AuxThm 2.4. Note that for $\lambda, \mu > -1/2$, we have

$$(x^\lambda, x^\mu)_{L_2(0,1)} = \frac{1}{\lambda + \mu + 1}.$$

Then the theorem follows from

$$G(x^{\lambda_0}, \dots, x^{\lambda_n}) = \frac{\prod_{n \geq i > k \geq 0} (\lambda_i - \lambda_k)^2}{\prod_{i=0}^n \prod_{k=0}^n (\lambda_i + \lambda_k + 1)},$$

and

$$G(x^r, x^{\lambda_0}, \dots, x^{\lambda_n}) = G(x^{\lambda_0}, \dots, x^{\lambda_n}) \cdot \frac{\prod_{k=0}^n (r - \lambda_k)^2}{(2r + 1) \prod_{k=0}^n (r + \lambda_k + 1)^2}.$$

□

Theorem (Part of [Theorem 1.1](#)). Let $\Lambda_\infty = \{0 = \lambda_0 < \lambda_1 < \dots < \lambda_n < \dots\}$ and $\lim_{n \rightarrow \infty} \lambda_n = \infty$. Then $\mathcal{M}(\Lambda_\infty)$ is dense in $L_2[0, 1]$ if and only if $\sum_{k=1}^\infty \lambda_k^{-1} = \infty$.

Remark 2.5. Let $a_k > -1$, the convergence or divergence of infinity product can be related to infinity sum:

- $\prod_k (1 + a_k)$ converges if and only if $\sum_k \log(1 + a_k)$ converges.

- $\prod_k(1+a_k)$ diverges to 0 (or $+\infty$) if and only if $\sum_k \log(1+a_k)$ diverges to $-\infty$ (or $+\infty$).

Proof. We note that the space of algebraic polynomials \mathbb{P} is dense in $L_2[0, 1]$. It is sufficient to show that

$$\forall r \in \mathbb{N} = \{0, 1, 2, \dots\}, \quad \lim_{n \rightarrow \infty} E(x^r, \Lambda_n)_2 = 0 \iff \sum_{k=1}^{\infty} \frac{1}{\lambda_k} = \infty.$$

" \Leftarrow " **Sufficiency.** Suppose that $\sum_{n=1}^{\infty} \lambda_n^{-1} = \infty$ and $r \in \mathbb{N} \setminus \Lambda_{\infty}$. Note that $0 \in \Lambda_{\infty}$, thus $r \geq 1$. There exists an index k_0 s.t. $\lambda_k > r$ whenever $k \geq k_0$. Then

$$\lim_{n \rightarrow \infty} E(x^r, \Lambda_n)_2 = \frac{1}{\sqrt{2r+1}} \frac{\prod_{k=0}^{\infty} |r - \lambda_k|}{\prod_{k=0}^{\infty} |r + \lambda_k + 1|} = C(r, k_0) \frac{\prod_{k=k_0}^{\infty} \left(1 - \frac{r}{\lambda_k}\right)}{\prod_{k=k_0}^{\infty} \left(1 + \frac{r+1}{\lambda_k}\right)},$$

where

$$C(r, k_0) = \frac{1}{\sqrt{2r+1}} \prod_{k=0}^{k_0-1} \frac{|r - \lambda_k|}{|r + \lambda_k + 1|}.$$

Denote

$$S_1 = \sum_{k=k_0}^{\infty} \log \left(1 - \frac{r}{\lambda_k}\right), \quad S_2 = \sum_{k=k_0}^{\infty} \log \left(1 + \frac{r+1}{\lambda_k}\right).$$

Then S_1 diverges to $-\infty$ (or the positive series $-S_1$ diverges to $+\infty$), if and only if the positive series

$$\sum_{k=k_0}^{\infty} \frac{r}{\lambda_k} = +\infty.$$

Similarly, S_2 diverges to ∞ if and only if the positive series

$$\sum_{k=k_0}^{\infty} \frac{r+1}{\lambda_k} = \infty.$$

Then $\lim_{n \rightarrow \infty} E(x^r, \Lambda_n)_2 = 0$ is obtained.

" \Rightarrow " **Necessity.** Otherwise, we suppose that $\sum_{k=1}^{\infty} \lambda_k^{-1} < \infty$. Then S_1 converges to a value ($\neq 0$), and S_2 converges to a value ($\neq 0$). Hence $\lim_{n \rightarrow \infty} E(x^r, \Lambda_n)_2 \neq 0$ leads to a contradiction. \square

Remark 2.6. The value $\lambda_0 = 0$ can be removed. In fact, let $\Lambda_{\infty} = \{0 < \lambda_1 < \dots < \lambda_n < \dots\}$ with $\lim_{n \rightarrow \infty} \lambda_n = +\infty$, then

$$\begin{aligned} \lim_{n \rightarrow \infty} E(1, \Lambda_n)_2 = 0 &\iff \prod_{k=1}^{\infty} \left(1 - \frac{1}{\lambda_k + 1}\right) = 0 \iff \sum_{k=1}^{\infty} \log \left(1 - \frac{1}{\lambda_k + 1}\right) = -\infty \\ &\iff \sum_{k=1}^{\infty} \frac{1}{\lambda_k + 1} = +\infty \iff \sum_{k=1}^{\infty} \frac{1}{\lambda_k} = +\infty. \end{aligned}$$

2.2.2 Case 2: $p = \infty$.

Our goal is to prove the (2.2) in [Lorentz (1996)], which is stated as following theorem:

AuxThm 2.6 (Theorem 5.5 in [Lorentz (1993)]). For $r > 0$, $\Lambda = \{\lambda_1, \lambda_2, \dots, \lambda_n\}$ with $\lambda_k > 0$, $k = 1, \dots, n$, we have

$$E(x^r, \Lambda)_{\infty} \leq \prod_{k=1}^n \frac{|r - \lambda_k|}{r + \lambda_k}. \quad (3)$$

Proof. For any $M > 0$ (it will be determined later), we put $\bar{r} = Mr$ and $\mu_k = M\lambda_k$. For any coefficients $c_k \in \mathbb{R}$, we set

$$b_k = \frac{\bar{r} + 1/2}{\mu_k + 1/2} c_k, \quad k = 1, 2, \dots, n,$$

and obtain

$$x^{\bar{r}+1/2} - \sum_{k=1}^n b_k x^{\mu_k+1/2} = \left(\bar{r} + \frac{1}{2}\right) \int_0^x \left[t^{\bar{r}-1/2} - \sum_{k=1}^n c_k t^{\mu_k-1/2} \right] dt. \quad (4)$$

Since $\mu_k - 1/2 > -1/2$, $k = 1, \dots, n$, by AuxThm 2.4 we can select c_k to satisfy

$$\left\| t^{\bar{r}-1/2} - \sum_{k=1}^n c_k t^{\mu_k-1/2} \right\|_{L^2(0,1)} = \frac{1}{\sqrt{2\bar{r}}} \prod_{k=1}^n \frac{|\bar{r} - \mu_k|}{\bar{r} + \mu_k}.$$

Then by Cauchy-Schwarz inequality and (4), we have $\forall x \in [0, 1]$ and $M > 0$

$$\left| x^{\bar{r}+1/2} - \sum_{k=1}^n b_k x^{\mu_k+1/2} \right| \leq \left(\bar{r} + \frac{1}{2}\right) \sqrt{x} \left\| t^{\bar{r}-1/2} - \sum_{k=1}^n c_k t^{\mu_k-1/2} \right\|_{L^2(0,1)},$$

which leads to

$$\left| x^{Mr} - \sum_{k=1}^n b_k x^{M\lambda_k} \right| \leq \frac{Mr + 1/2}{\sqrt{2Mr}} \prod_{k=1}^n \frac{|r - \lambda_k|}{r + \lambda_k}. \quad (5)$$

By choosing $M = 1/(2r)$ and taking the transform $u = x^{1/(2r)}$ on (5), we have $\forall u \in [0, 1]$

$$\left| u^r - \sum_{k=1}^n b_k u^{\lambda_k} \right| \leq \prod_{k=1}^n \frac{|r - \lambda_k|}{r + \lambda_k},$$

which give rise to (3). □

Theorem (Part of Theorem 1.1). *Let $\Lambda_\infty = \{0 = \lambda_0 < \lambda_1 < \dots < \lambda_n < \dots\}$ and $\lim_{n \rightarrow \infty} \lambda_n = \infty$. Then $\mathcal{M}(\Lambda_\infty)$ is dense in $C[0, 1]$ if and only if $\sum_{k=1}^\infty \lambda_k^{-1} = \infty$.*

Remark 2.7. $\lambda_0 = 0$ must be included in Λ_∞ .

Proof. We note that \mathbb{P} is dense in $C[0, 1]$. It is sufficient to show that

$$\forall r \in \mathbb{N} = \{0, 1, 2, \dots\}, \quad \lim_{n \rightarrow \infty} E(x^r, \Lambda_n)_\infty = 0 \iff \sum_{k=1}^\infty \frac{1}{\lambda_k} = \infty.$$

" \Leftarrow " **Sufficiency.** Suppose that $\sum_{n=1}^\infty \lambda_n^{-1} = \infty$ and $r \in \mathbb{N} \setminus \Lambda_\infty$. Note that $0 \in \Lambda_\infty$, thus $r \geq 1$. There exists an index k_0 s.t. $\lambda_k > r$ whenever $k \geq k_0$. Then

$$\lim_{n \rightarrow \infty} E(x^r, \Lambda_n)_\infty \leq \prod_{k=1}^\infty \frac{|r - \lambda_k|}{r + \lambda_k} = C(r, k_0) \frac{\prod_{k=k_0}^\infty \left(1 - \frac{r}{\lambda_k}\right)}{\prod_{k=k_0}^\infty \left(1 + \frac{r}{\lambda_k}\right)}, \quad \text{where } C(r, k_0) = \prod_{k=1}^{k_0-1} \frac{|r - \lambda_k|}{r + \lambda_k}.$$

Denote

$$S_1 = \sum_{k=k_0}^\infty \log \left(1 - \frac{r}{\lambda_k}\right), \quad S_2 = \sum_{k=k_0}^\infty \log \left(1 + \frac{r}{\lambda_k}\right).$$

Then S_1 diverges to $-\infty$ and S_2 diverges to $+\infty$, leading to $\lim_{n \rightarrow \infty} E(x^r, \Lambda_n)_\infty = 0$.

" \Rightarrow " **Necessity.** Note that

$$E(x^r, \Lambda_n)_\infty \geq E(x^r, \Lambda_n)_2.$$

Then $\forall r \in \mathbb{N}$, $\lim_{n \rightarrow \infty} E(x^r, \Lambda_n)_\infty = 0$ gives rise to $\lim_{n \rightarrow \infty} E(x^r, \Lambda_n)_2 = 0$, which leads to $\sum_{k=1}^\infty \lambda_k^{-1} = \infty$. □

2.2.3 Case 3: $2 < p < \infty$.

Our goal is to prove the [Theorem 2.2](#) in [\[Lorentz \(1996\)\]](#):

Theorem 2.2 ([Theorem 2.2](#) in [\[Lorentz \(1996\)\]](#)). *Let $2 < p < \infty$ and $\Lambda = \{\lambda_0, \lambda_1, \dots, \lambda_n\}$ with distinct elements and $\lambda_k > -1/p$. For any $r > -\frac{1}{p}$, we have*

$$E(x^r, \Lambda)_p \leq \frac{1 + 1/p}{(2r + 2/p)^{1/p}} \prod_{k=0}^n \frac{|r - \lambda_k|}{r + \lambda_k + 2/p}. \quad (2.4)$$

Lemma 2.1 ([Lemma 2.1](#) in [\[Lorentz \(1996\)\]](#)). *Let $1 \leq q < p < \infty$ and let $-\frac{1}{q} < \ell_0 < \ell_1 < \dots < \ell_n$. For arbitrary real numbers a_0, a_1, \dots, a_n and*

$$b_k := \frac{1 + \ell_k + \frac{1}{p}}{1 + \frac{1}{p}} a_k, \quad 0 \leq k \leq n,$$

we have

$$\left\| x^{\frac{1}{q} - \frac{1}{p}} - \sum_{k=0}^n a_k x^{\ell_k + \frac{1}{q} - \frac{1}{p}} \right\|_p \leq \left(1 + \frac{1}{p} \right) \left\| 1 - \sum_{k=0}^n b_k x^{\ell_k} \right\|_q. \quad (2.3)$$

Proof. Let us denote $K := 1 + \frac{1}{p}$ and for $0 < x \leq 1$

$$\begin{aligned} Q(x) &:= \sum_{k=0}^n b_k x^{\ell_k}, \quad h(x) := x^{\frac{1}{p}} (1 - Q(x)), \\ g(x) &:= K x^{\frac{1}{q} - 1 - \frac{2}{p}} \int_0^x h(t) dt. \end{aligned}$$

One easily verifies that g is the function on the left hand side of [\(2.3\)](#). Our goal is to show

$$\|g\|_p \leq K \|1 - Q(x)\|_q.$$

To achieve this goal, we employ Hölder type inequalities.

Hölder inequality: For any $1 \leq p, q \leq \infty$ that satisfy $\frac{1}{p} + \frac{1}{q} = 1$. Suppose that $f \in L_p(\Omega)$, $g \in L_q(\Omega)$ and $fg \in L_1(\Omega)$, then

$$\|fg\|_{L_1(\Omega)} \leq \|f\|_{L_p(\Omega)} \|g\|_{L_q(\Omega)}.$$

Firstly, by Hölder's inequality, we have for $0 < x \leq 1$,

$$\begin{aligned} |g(x)| &\leq K x^{\frac{1}{q} - 1 - \frac{2}{p}} \int_0^x |h(t)| dt \leq K x^{\frac{1}{q} - 1 - \frac{2}{p}} \left(\int_0^x 1 dt \right)^{1 - \frac{1}{q}} \left(\int_0^x |h(t)|^q dt \right)^{\frac{1}{q}} \\ &= K x^{-\frac{2}{p}} \left(\int_0^x |h(t)|^q dt \right)^{\frac{1}{q}} =: K \left(\int_0^1 F(x, t) dt \right)^{\frac{1}{q}}, \end{aligned}$$

where

$$F(x, t) := \begin{cases} x^{-\frac{2q}{p}} |h(t)|^q, & \text{if } 0 \leq t < x, \\ 0, & \text{otherwise.} \end{cases} \quad \text{Note that } F(x, t) \in [0, 1] \times [0, 1].$$

Hölder-Minkowski inequality (see [\[Bahouri \(2011\), p.4\]](#)) states:

Let (X_1, μ_1) and (X_2, μ_2) be two measure spaces and f be a nonnegative measurable function over $X_1 \times X_2$. For all $1 \leq q \leq p \leq \infty$, we have

$$\left\| \|f(x_1, \cdot)\|_{L_q(X_2, \mu_2)} \right\|_{L_p(X_1, \mu_1)} \leq \left\| \|f(\cdot, x_2)\|_{L_p(X_1, \mu_1)} \right\|_{L_q(X_2, \mu_2)}.$$

Then by Hölder-Minkowski inequality, we have

$$\begin{aligned}
\|g\|_p &\leq K \left[\int_0^1 \left(\int_0^1 F(x, t) dt \right)^{\frac{p}{q}} dx \right]^{\frac{1}{p}} = K \left\| \|F(x, \cdot)^{\frac{1}{q}}\|_q \right\|_p \\
&\leq K \left\| \|F(\cdot, t)^{\frac{1}{q}}\|_p \right\|_q = K \left[\int_0^1 \left(\int_0^1 F(x, t)^{\frac{p}{q}} dx \right)^{\frac{q}{p}} dt \right]^{\frac{1}{q}} \\
&= K \left[\int_0^1 \left(\int_t^1 x^{-2} |h(t)|^p dx \right)^{\frac{q}{p}} dt \right]^{\frac{1}{q}} \\
&= K \left[\int_0^1 |h(t)|^q \left(\frac{1}{t} - 1 \right)^{\frac{q}{p}} dt \right]^{\frac{1}{q}} \leq K \left[\int_0^1 |h(t)|^q \left(\frac{1}{t} \right)^{\frac{q}{p}} dt \right]^{\frac{1}{q}} \\
&= K \left[\int_0^1 |t^{-\frac{1}{p}} h(t)|^q dt \right]^{\frac{1}{q}} = K \left[\int_0^1 |1 - Q(t)|^q dt \right]^{\frac{1}{q}}.
\end{aligned}$$

where the inequality holds since $0 < \frac{q}{p} < 1$ and $0 < \frac{1}{t} - 1 < \infty$ and

$$\left(\frac{1}{t} - 1 \right)^{\frac{q}{p}} \leq \left(\frac{1}{t} \right)^{\frac{q}{p}}.$$

□

Remark 2.8. Note that $q < p$, then Lemma 2.1 is some kind of Inverse Inequality: Higher regularity norm bounded by lower regularity norm.

Proof of Theorem 2.2. To prove (2.4), our goal is to employ Lemma 2.1 and construct the formula

$$E(x^r, \Lambda)_p \leq \left\| x^{\frac{1}{2} - \frac{1}{p}} - \sum_{k=0}^n a_k x^{\ell_k + \frac{1}{2} - \frac{1}{p}} \right\|_p, \quad \text{for some } \ell_k > -\frac{1}{2} \text{ and } a_k.$$

To achieve this, for any a_k , $0 \leq k \leq n$, which will be determined later, we have

$$E(x^r, \Lambda)_p \leq \|x^r - \sum_{k=0}^n a_k x^{\lambda_k}\|_p = \left[\int_0^1 \left(x^r - \sum_{k=0}^n a_k x^{\lambda_k} \right)^p dx \right]^{\frac{1}{p}}.$$

By a variable transform $x = y^\rho$, $\rho > 0$, which is invariant under the interval $[0, 1]$ and ρ will be determined later, we have

$$\begin{aligned}
E(x^r, \Lambda)_p &\leq \left[\int_0^1 \left(y^{\rho r} - \sum_{k=0}^n a_k y^{\rho \lambda_k} \right)^p \rho y^{\rho-1} dy \right]^{\frac{1}{p}} \\
&= \rho^{\frac{1}{p}} \left\| y^{\rho r + \frac{\rho}{p} - \frac{1}{p}} - \sum_{k=0}^n a_k y^{\rho \lambda_k + \frac{\rho}{p} - \frac{1}{p}} \right\|_p.
\end{aligned}$$

Let $\rho r + \frac{\rho}{p} = \frac{1}{2}$, we obtain $\rho = \frac{p}{2(pr+1)}$. Let $\ell_k = \frac{p(\lambda_k - r)}{2(pr+1)} > -\frac{1}{2}$, it is easy to examine that $\ell_k + 1/2 > 0$, by Lemma 2.1, we have

$$\begin{aligned}
E(x^r, \Lambda)_p &\leq \left(\frac{p}{2(pr+1)} \right)^{\frac{1}{p}} \left\| y^{\frac{1}{2} - \frac{1}{p}} - \sum_{k=0}^n a_k y^{\ell_k + \frac{1}{2} - \frac{1}{p}} \right\|_p \\
&\leq \left(\frac{p}{2(pr+1)} \right)^{\frac{1}{p}} \left(1 + \frac{1}{p} \right) \left\| 1 - \sum_{k=0}^n b_k y^{\ell_k} \right\|_2.
\end{aligned} \tag{6}$$

Since a_k is arbitrary, hence b_k is also arbitrary. Take the infimum on the right hand side of (6) over b_k , and by Theorem 2.4, we have

$$E(x^r, \Lambda)_p \leq \left(\frac{p}{2(pr+1)} \right)^{\frac{1}{p}} \left(1 + \frac{1}{p} \right) \prod_{k=0}^n \frac{|\ell_k|}{\ell_k + 1} = \frac{1 + 1/p}{(2r + 2/p)^{1/p}} \prod_{k=0}^n \frac{|r - \lambda_k|}{r + \lambda_k + 2/p}.$$

□

Theorem (Part of Theorem 1.1). *Let $\Lambda_\infty = \{0 = \lambda_0 < \lambda_1 < \dots < \lambda_n < \dots\}$ and $\lim_{n \rightarrow \infty} \lambda_n = \infty$. Then $\mathcal{M}(\Lambda_\infty)$ is dense in $L_p[0, 1]$, $2 < p < \infty$, if and only if $\sum_{k=1}^{\infty} \lambda_k^{-1} = \infty$.*

Proof. We note that \mathbb{P} is dense in $L_p[0, 1]$. It is sufficient to show that

$$\forall r \in \mathbb{N} = \{0, 1, 2, \dots\}, \quad \lim_{n \rightarrow \infty} E(x^r, \Lambda_n)_p = 0 \iff \sum_{k=1}^{\infty} \frac{1}{\lambda_k} = \infty.$$

" \Leftarrow " **Sufficiency.** Suppose that $\sum_{n=1}^{\infty} \lambda_n^{-1} = \infty$ and $r \in \mathbb{N} \setminus \Lambda_\infty$. Note that $0 \in \Lambda_\infty$, thus $r \geq 1$. There exists an index k_0 s.t. $\lambda_k > r$ whenever $k \geq k_0$. Then

$$\lim_{n \rightarrow \infty} E(x^r, \Lambda_n)_p \leq \frac{1 + 1/p}{(2r + 2/p)^{1/p}} \frac{\prod_{k=0}^{\infty} |r - \lambda_k|}{\prod_{k=0}^{\infty} |r + \lambda_k + 2/p|} = C(r, k_0) \frac{\prod_{k=k_0}^{\infty} \left(1 - \frac{r}{\lambda_k}\right)}{\prod_{k=k_0}^{\infty} \left(1 + \frac{r+2/p}{\lambda_k}\right)},$$

where

$$C(r, k_0) = \frac{1 + 1/p}{(2r + 2/p)^{1/p}} \prod_{k=0}^{k_0-1} \frac{|r - \lambda_k|}{|r + \lambda_k + 2/p|}.$$

Denote

$$S_1 = \sum_{k=k_0}^{\infty} \log \left(1 - \frac{r}{\lambda_k} \right), \quad S_2 = \sum_{k=k_0}^{\infty} \log \left(1 + \frac{r+2/p}{\lambda_k} \right).$$

Then S_1 diverges to $-\infty$ and S_2 diverges to $+\infty$, leading to obtain $\lim_{n \rightarrow \infty} E(x^r, \Lambda_n)_p = 0$.

" \Rightarrow " **Necessity.** Note that

$$E(x^r, \Lambda_n)_p \geq E(x^r, \Lambda_n)_2.$$

Then $\forall r \in \mathbb{N}$, $\lim_{n \rightarrow \infty} E(x^r, \Lambda_n)_p = 0$ gives rise to $\lim_{n \rightarrow \infty} E(x^r, \Lambda_n)_2 = 0$, which leads to $\sum_{k=1}^{\infty} \lambda_k^{-1} = \infty$. □

Remark 2.9. *The value $\lambda_0 = 0$ can be removed. In fact, let $\Lambda_\infty = \{0 < \lambda_1 < \dots < \lambda_n < \dots\}$ with $\lim_{n \rightarrow \infty} \lambda_n = +\infty$,*

$$\begin{aligned} \prod_{k=1}^{\infty} \left(1 - \frac{2/p}{\lambda_k + 2/p} \right) = 0 &\iff \sum_{k=1}^{\infty} \log \left(1 - \frac{2/p}{\lambda_k + 2/p} \right) = -\infty \\ &\iff \sum_{k=1}^{\infty} \frac{2/p}{\lambda_k + 2/p} = +\infty \iff \sum_{k=1}^{\infty} \frac{1}{\lambda_k} = +\infty. \end{aligned}$$

Then if $\sum_{k=1}^{\infty} \lambda_k^{-1} = \infty$, we have $\lim_{n \rightarrow \infty} E(1, \Lambda_n)_p = 0$. Conversely, if $\lim_{n \rightarrow \infty} E(1, \Lambda_n)_p = 0$, which leads to $\lim_{n \rightarrow \infty} E(1, \Lambda_n)_2 = 0$, we have $\sum_{k=1}^{\infty} \lambda_k^{-1} = \infty$.

2.2.4 Conclusion and remarks.

Let $2 \leq p \leq \infty$ and $\Lambda = \{\lambda_k\}_{k=0}^n$ with $\lambda_k > -1/p$. Then for any $r > -1/p$, we have

$$E(x^r, \Lambda)_p \leq \frac{1 + 1/p}{(2r + 2/p)^{1/p}} \prod_{k=0}^n \frac{|r - \lambda_k|}{r + \lambda_k + 2/p}.$$

A Variant of Dense Property.

(see also Section 5 of Chapter 11 in [\[Lorentz \(1993\)\]](#))

Let $C[0, +\infty]$ be the space of continuous functions on $[0, +\infty]$, which have a finite limit for $t \rightarrow \infty$. Exponential sums

$$\sum_{k=0}^n a_k e^{-\lambda_k t}$$

approximate arbitrarily closely each function $f \in C[0, +\infty]$ if and only if $\lambda_0 = 0$, $\lambda_k > 0$ for $k \geq 1$ and $\sum_{k=1}^{\infty} \lambda_k^{-1} = \infty$.

References

- [Lorentz (1996)] Lorentz G G, von Golitschek M, Makovoz Y. Constructive approximation: advanced problems[M]. Berlin: Springer, 1996.
- [Erdélyi (2001)] Erdélyi T, Johnson W B. The “Full Müntz Theorem” in $L_p[0, 1]$ for $0 < p < \infty$ [J]. Journal D’Analyse Mathématique, 2001, 84(1): 145-172.
- [Almira (2007)] Almira J M. Muntz type theorems I[J]. arXiv preprint arXiv:0710.3570, 2007.
- [Agler (2022)] Agler J, McCarthy J. Asymptotic Müntz-Szász Theorems[J]. arXiv preprint arXiv:2206.12487, 2022.
- [Wang Renhong (1999)] Wang Renhong. Numerical Approximation (in Chinese)[M]. Higher Education Press, 1999.
- [Lorentz (1993)] DeVore R A, Lorentz G G. Constructive Approximation[M]. Springer-Verlag, 1993.
- [Bahouri (2011)] Bahouri H. Fourier Analysis and Nonlinear Partial Differential Equations[M]. Grundlehren der Mathematischen Wissenschaften, 2011, 343.
- [Adams (2003)] Adams R A, Fournier J J F. Sobolev spaces[M]. Elsevier, 2003.
- [Borwein (1995)] Borwein P, Erdelyi T. Polynomials and Polynomial Inequalities[M]. Springer Science & Business Media, 1995.
- [Powell (1981)] Powell M J D. Approximation theory and methods[M]. Cambridge university press, 1981.
- [Brezis (2011)] Brezis H. Functional analysis, Sobolev spaces and partial differential equations[M]. New York: Springer, 2011.